November 9, 2024

Dear Erzo,

Thank you very much for your feedback on our paper "Rationing Medicine Through Bureaucracy: Authorization Restrictions in Medicare" as part of the evaluation process at the *AER*.

We want to note up front that we greatly appreciate the time and effort that you, John Friedman, and the four reviewers put into evaluating the paper. We were very happy to see that all of you were as excited about the topic as we are. We have taken the many comments we received to heart, and we believe the new version we are resubmitting is greatly improved as a result. We also want to apologize for the extended time it took to resubmit. The last year has brought many expected and unexpected productivity shocks to our group, including multiple family-member cancer diagnoses, multiple job searches, multiple babies, and a wedding, that led to the significant delay in getting this back to you.

In this document we will start by providing a short summary of the changes we made. We will then respond to your decision letter. We will then provide detailed responses to the reports from each of the reviewers individually. As you will notice, this document ended up being quite long. To make it easier to navigate, click the hyperlinks below to be taken to a specific section of the document. The hyperlinks are also embedded in the header of the subsequent pages. In our experience as reviewers, we have noticed that the *AER* formatting system sometimes destroys hyperlinks in PDF files; if this happens, you can find another version of this document at `https://zarekcb.github.io/PriorAuth_ResponseLetter.pdf`.

Thank you again for considering the paper for publication in the *AER*. We are excited about this revised version and look forward to hearing your thoughts on it. Please let us know if there are any questions or concerns in your consideration of this resubmission.

Best regards,

Zarek Brot
Samantha Burn
Timothy Layton
Boris Vabson

Summary of Changes

Exposition. We have restructured the paper slightly, in response to confusion by all reviewers, but particularly a comment by Reviewer 4 (4.2). We have removed the section previously called "Substitution Patterns and Spending Effects." In that section, we (1) estimated a structural model of demand; and (2) used it to simulate how demand and spending would change in the complete absence of prior authorization. The Reviewer made the convincing point that there was no reason this should be a separate section.

We have now taken the analysis of substitution patterns and folded it into the section with our main analysis on direct effects, and have changed the focus of this towards linear regression (see below). This section now reports all key results, which all come from a single empirical design and model. Section 5 of the paper now focuses entirely on the counterfactual simulation of banning prior authorization and the empirical analysis of the trade-off between effects of such a ban on drug spending and administrative costs. This section does include the structural demand model from the prior version of the paper, as we argue that it is necessary for this counterfactual simulation of turning prior authorization restrictions off for *all* drugs (versus just for a single drug). We hope the paper now reads more cleanly than it did before.

We have also expanded our substantive discussion of identification and limits to identification given our instrumental variation, as well as how we discuss aggregation across drugs. We have overhauled (and merged) our research design and first stage subsections (see E.2h, E.2i, E.5). We have also added language to discuss the origins of the variation in the use of prior authorization (R1.1, R1.3, R2.1, R3.MC1). We have taken out the case study on oral anticoagulants (E.4, R2.3, R2.4).

We have rewritten the consumer surplus section substantially in response to multiple comments that it was difficult to follow (E.6, R2.7, R4.3, R4.4), and we modified the analyses in that section to use linear demand curves as our main specification instead of constant semi-elasticity demand curves. We hope the new version is clearer. Finally, we have taken great care to tone down the use of normative language that the editor and some reviewers felt was not justified by our analysis (E.8e, R1.1, R3.1).

Analysis. The editor as well as multiple reviewers expressed confusion about apparent inconsistencies between the multiple empirical specifications in the original manuscript (E.2, R3.3a, R4.2, R4.4). The editor's suggestion was to remove *all* nonlinear modeling from the paper (i.e., the nested logit demand system, and the Poisson regression approach for estimating WTP). He instead suggested that we take a linear approach to both.

In response, we have indeed estimated a linear model of substitution responses to prior authorization, in the new Section 3.4 ("Substitution Responses"). Reassuringly, the results from this section are of very similar magnitudes to those we obtained using the logit estimation. This also enables us to more clearly present heterogeneity in diversion (E.2a, R2.2).

We also made the main specification in the revealed preference section a linear specification. We

have altered it to hew closer to the method used by Gross et al. (2022), from which we borrowed the design. We include the prior Poisson approach as a robustness check. We have also included a robustness check here to deal with both editors' points about aggregation (E.3, E.8f). We note that these results have changed in magnitude—the linear approach has losses in WTP that are larger than our prior estimates by a factor of 3-4, meaning that, in some (though not all) specifications, the consumer surplus losses are larger than the net financial savings, which was not previously true. We have edited the text to reflect this change in our results.

In contrast to the suggestion, we have *not* removed the logit analysis completely. We did understand the editor and reviewers' confusion, however, and we made some significant changes to exposition in order to make the contribution of the structural model clearer. In the revision, we have reduced its presence substantially to *only* the section where we think it is absolutely essential (comparing spending effects to admin costs in the counterfactual simulation banning all prior authorization restrictions), and have included a much more explicit justification of its presence than existed in the original manuscript. We think that this choice strikes a balance between (1) the view of the editor and reviewers that the paper would benefit from being more straightforward and (2) the complexity of thinking about overall effects of the relevant policy, a complete ban on prior authorization restrictions, in a world with (difficult-to-estimate) drug-to-drug substitution patterns.

Changes to figures and tables.

- We now include two new figures that plot, for a subset of drugs, a simple comparison in utilization between beneficiaries assigned to plans that restrict those drugs versus those assigned to plans that do not, as suggested by the editor (E.2a). These now appear in Figure 5a and 5b, and they exhibit our results quite clearly, so we are very grateful for this suggestion.
- Our main regression table (Table 5) is no longer a waterfall table of multiple specifications for the same outcome. It now has multiple utilization outcomes, as suggested by E.2c.i. We have moved the prior table to Appendix Table A6.
- We now include a new figure reporting heterogeneity in the diversion ratio by beneficiaries (R2.2). This appears in Figure 7.
- We have changed the aggregate analysis of health effects to use IV rather than ITT, to retain consistency with the rest of the paper (E.2i) and have changed the regressor to reflect continuous variation in exposure to prior authorization (E.4). We now find even less conclusive results–our point estimates now suggest that prior authorization *reduces* mortality, though again with large standard errors.

Other. Zarek is in the process of legally changing his last name from "Brot-Goldberg" to "Brot" after getting married and is transitioning to that name professionally. It has been updated in the manuscript, but we wanted to flag it in case there was any confusion.

Editor's Decision Letter

(E.1) *I didn't understand why you needed to limit sample to people being randomized into a plan because their previous plans no longer qualifying as a benchmark plan. You say you do this in order to get access to past prescriptions, but you don't use past prescriptions in your main specifications. I would therefore use all people who are randomized into a plan, for whatever reason. Please be clear in the paper that your estimates apply to the selected sample of those who failed to make an active choice.*

Thanks for this comment. When we were implementing this recommendation, we realized that we were already including more than just those whose plans lost benchmark status. Our sample also already included those whose plans exited from the market altogether. We apologize for the confusion.

Based on your suggestion, we also considered including those who were randomized at the time of entry into Medicare. In the end, we decided not to include them, for two reasons. First, we do not observe prior utilization for this group, so testing balance or using them for the health effects analysis is not possible. Even more importantly, however, we typically do not observe a full year of post-randomization data for them with continuous enrollment in the same plan, as they typically enroll in the month they turn 65 rather than in January, and enroll in a new plan in the following January. This would require some sort of statistical adjustment for this shortened window. In the spirit of what we perceived as your high-level goal of having us make this paper be more transparent and unified (E.2), we felt this would not accomplish this goal.

We clarify all of this in the new version of the paper. We've pasted the modified paragraphs below for convenience.

> *We additionally restrict to two groups of LIS beneficiaries who faced the automatic reassignment mechanism described in Section 2.1: (1) those who were previously automatically-enrolled in a benchmark plan, whose plan subsequently lost benchmark status by charging a monthly premium above the premium subsidy and (2) those whose prior plan exited the market entirely. We focus on these beneficiaries, rather than new Medicare enrollees, for two key reasons: (1) we observe a full year of post-randomization data for them, and (2) for this group we can observe pre-assignment data, providing a useful outcome for placebo tests and useful information about the set of drugs demanded by the beneficiary that we use in some analyses. We exclude beneficiaries whose reassignment is expected to be non-random based on program rules.*

(E.2) *Transparent and unified specifications (Reinforcing R4 comment 2)*
What is so nice about your setting is that you have many natural experiments: 8 years times

N service regions (BTW, what is N? Roughly 27 or so?). This should allow you to adopt a transparent estimation approach, which remains the same or very similar in all sections of the paper, doesn't take 3 days to estimate, doesn't require the imposition of structure (or produce lambdas outside the required range of 0-1), doesn't require highly specialized procedures to produce standard errors, doesn't require the factually incorrect assumption that a person consumes at most one specific drug within a therapeutic class, and doesn't require assuming that usage and spending are proportional. In short, please drop the nested logit specification and the simulations that go with it.

a. ***NOTE: Discussed below, see E.2a***
b. *I would strongly suggest using linear models that have coefficients that are directly interpretable. This limits the amount of structure you impose. It will also help convey to the reader what you are doing.*
c. *I think Tables 5 and 6 could be estimated with the same basic model, as laid out in your equation 1. I see 6 relevant types of outcome variables:*
    i. *Effects on spending and usage of the drug in question. Usage is as in your current table 5, but you would also add spending on the drug in question.*
    ii. *Effects on total spending and total usage on all other drugs, except the drug in question, that belong to the same therapeutic class. This in effect assumes no cross-therapeutic-class substitution, as you also do with your nested logit. Note that one metric of total usage could be a dummy for no usage of any other drug in this class.*
    iii. *Effects on total spending and total usage on all other drugs, except the drug in question (so not conditioning on therapeutic class). This imposes no substitution restrictions, but I expect that these estimates might be too noisy. In the latter case, just put them in the response letter, but leave them out of the paper (and say in the paper why you don't do this).*
d. *The benefit of this specification (at least i and ii) is that you can also make the estimates specific within therapeutic class, and report the heterogeneity.*
e. *It might be interesting to do the above table both in levels and in percentages of base spending or usage of the drug in question. For the percentage effects, you would divide all the spending variables by the mean spending for the drug among those for whom the drug is unrestricted and you would divide all the usage variable by the mean usage among those for whom the drug is unrestricted.*

Thank you for this comment. We spent a lot of time on it, as well as your comment E.2a below. First, in the spirit of your comment c.i., we have replaced Table 5 with a new table that estimates our main specification with three outcomes: First, whether the beneficiary ever used the drug (our primary outcome from the initial draft); second, the total number of drug days supplied; and third, spending (net of rebates). We have moved the prior table to Appendix Table A6.

We have followed your primary suggestion. We now include transparent linear regression specifications where we estimate the effect of prior auth on drug A on (1) use of every other drug in the same class as A, *excluding* A, which is your suggestion (ii) above; and (2) use of every drug in the same class as A, *including* A. This new analysis now appears in Table 6. We find results that are consistent in magnitude with our prior logit results: For both our "ever used" and "days" measure of quantities, roughly half of use deterred from the focal drug is made up for by use of other drugs. In contrast, 8% of dollars not spent on the focal drug due to prior auth are diverted to a substitute drug. This is all discussed in the new Section 3.4, which is dedicated to estimating substitution effects.

We did not use your suggestion (iii) in the main text because, as you anticipated, it produces very odd estimates. We have reproduced a table for it below. We did not include a regression where the outcome is "ever used" for all other drugs because there is essentially no variation in that outcome—virtually everyone in our sample is taking at least one drug.

Table E1. Regression for All Other Drugs

	Days Supply	Spending
$\text{Auth}^{\text{Enrolled}}$	-2.950	-36.152
	(1.652)	(5.896)
Auth^{Sub}	9.426	-67.609
	(6.588)	(20.880)
PA % Effect	-0.2	-1.1
Control Mean	1518.123	3172.035
Reweighted Control Mean	1520.012	3194.632

In the spirit of your comments (d) and (e), we have produced, for each heterogeneity analysis where we create a figure of coefficient estimates for the direct effect, a corresponding figure for substitution to other drugs in class. We find the coefficient a bit tricky to interpret both in levels and in percent changes, so we have instead plotted it in terms of the estimated diversion ratio, which is the answer to the question "for every use/dollar of drug A reduced by prior authorization on drug A, how much does use/spending of other drugs increase?" We think this is the best measure to compare these estimates against the direct effect estimates, which we interpreted as the spirit of this suggestion. A version with beneficiary heterogeneity is now given in Figure 7. While we did estimate this for other sources of heterogeneity, our power drops significantly and the effects vary wildly (in the way they do in our Appendix Figures when we cut by e.g. class or by price deciles). We didn't feel these were informative enough to put in the main text.

One major wrinkle with this approach: In contrast to our main regressions, in these regressions, plan fixed effects are no longer well-identified. This took us an embarrassingly long time to figure out, leading in part to the delay in resubmitting. To understand this, it is worth remembering that the plan fixed effects in our specification are identified by comparing use of drugs that are *never* restricted across plans. However, this is essentially the exact same thing that your suggested regressions are try-

ing to estimate: differences in the use of never-restricted drugs between plans that do or do not restrict a substitute. In other words, the plan fixed effects are only identified when there are drugs that are never themselves restricted *and* that never have substitutes that are restricted. This essentially never happens in our setting, as our best way of identifying substitute drugs is via therapeutic classes and every class with nontrivial use has at least one restricted drug. Technically, the plan effects can still be identified by leveraging variation in the percent of substitutes that are restricted to extrapolate to what use would be if that percent were zero, but this requires significant out-of-sample extrapolations that can lead to strange and misleading results (e.g. wrong-signed effects). Further, when our outcome is "the use of other drugs" and the focal drug in question is a never-restricted drug, the outcome in question *includes the utilization of sometimes-restricted drugs*, making the plan fixed effects confusing to interpret. We therefore omit plan fixed effects in these regressions.

We interpreted the spirit of your comment (and the reviewers', in e.g. R3.3a, R4.2) as making sure we have a unified specification for as much of the paper as we can. In that spirit, we have changed our primary specification for the focal drug regressions to also omit plan FEs, though we include them in a robustness check in Appendix Table A7, where we show that when the outcome is of the focal drug the inclusion of plan effects has little to no effect on our estimates. We have thus updated **all** of the focal drug analyses to drop plan FEs. There have been no substantive changes to any of our conclusions, consistent with the idea that we are able to observe and control for essentially all differences across plans (prior auth and exclusion on the focal drug, prior auth and exclusion on substitute drugs).

After presenting main results for our linear specification estimating effects of prior authorization restrictions on use of restricted drugs and substitutes, we then proceed with a section that analyzes the trade-off between spending reductions and administrative costs associated with prior authorization via a counterfactual simulation banning all prior authorization restrictions.[1] For this section, we need to perform a counterfactual simulation where we remove prior authorization restrictions not just for a single drug but for *all* drugs. This simulation is important, as it captures the effects of not just turning off prior authorization for a given drug but also for all of the *substitutes* for that drug. The latter is the true key policy parameter when assessing whether prior authorization should be banned, as many lobbying groups and health policy advocates and researchers have suggested. To perform this counterfactual, we need to estimate a full demand system for drugs so that we can capture substitution patterns not just from turning prior authorization restrictions on one drug off but also from turning prior authorization restrictions on the substitutes for that drug off, especially when we believe there to be nonlinearities in such effects (as we would certainly expect here). Indeed, an entire literature in IO exists that shows the utility of using discrete choice models to estimate flexible substitution patterns (see e.g. Hausman, Leonard, and Zona 1994 and Nevo 2011), and we rely on that literature to show the need for adding structure to the problem in order to perform this set of counterfactuals.

We explain all of this at the start of Section 4. We include the explanation here for convenience:

[1] As mentioned in the summary, this has now been compressed into a single section. R4.2 helped us understand why this previously read in a confusing way.

> *Importantly, simply scaling up our reduced form estimates may fail to account for interactions between restrictions on drugs. For instance, restricting drugs A and B is likely to have different effects than the sum of the effects of restricting each drug individually. This is because restricting A may induce substitution to B (and vice versa). These interactions may be important for understanding the full extent of the effects of allowing prior authorization as an institution. Further, turning restrictions off for one drug at a time (the effect captured in our reduced form approach) may induce substitution to another restricted drug, thus not reducing administrative costs. As discussed earlier, we lack the variation to cleanly identify such rich substitution patterns, making this empirically difficult without imposing additional structure.*
>
> *Thus, we proceed by explicitly modeling demand for drugs using a standard micro-founded model of demand for restricted and unrestricted drugs (and their substitutes) based on utility maximization. We then use the estimated model to simulate removing all prior authorization restrictions and measure the effects on utilization, spending, and administrative costs. Such an approach requires us to make additional data restrictions and modeling assumptions relative to our reduced-form approach, but we show that these choices have relatively minimal effects via comparisons to our reduced-form estimates.*

Critically, we now use the nested logit analysis *only* for this counterfactual, and, in response to you and many of the reviewers (e.g. R4.2), we now explain this choice more thoroughly and directly.

One additional comment related to the choice to retain the structural model for the counterfactual of removing prior authorization from *all* drugs is that, even if our goal was to estimate the average effect of restricting an *individual* drug (as you suggest in E.8d), it would have been challenging to measure effects on administrative costs: some substitution due to prior authorization would have been to a different restricted drug; we would thus need to estimate the inframarginal use of the restricted drug *and* marginal use of substitute restricted drugs. This would necessitate identifying substitution patterns between specific drug pairs; as we now discuss in Section 3.4, we lack the instrumental variation to do so without the structure we add in Section 5.

One last thing: We now note (on page 9) that there are 34 service regions.

(E.2a) *In principle, you can show the effects of prior authorizations graphically by a comparison of means for any given drug – this is the beauty of having randomization! The only control that you truly need is the year*market FE, which are your randomization strata. For each drug, you can plot the mean usage (or spending) on that drug for consumers randomly assigned to plans with prior authorizations and for those assigned to plans without, and the difference between the two, of course controlling for the year*market FEs. You exclude from this plot consumers placed in plans where the drug is excluded. It would be nice to show this graphically for the most common drugs for which there is meaningful variation in prior authorization. You can do a*

*similar plot for total usage (or spending) on other drugs in the same therapeutic class. To select for which drugs you show this graph, I would suggest you calculate for each drug and market: minshare=minimum(fraction of plans in that market that require prior authorizations this drug, fraction of plans in that market that don't require prior authorizations for this drug). You then sum for each drug across markets: (number of users of this drug in that market)*minshare. Finally, you select (say) the 30 drugs for which this sum is highest. This would be a systematic way of finding the most commonly used drugs for which you have meaningful variation within markets in prior authorization. Starting off with simple graphical evidence is a great way to explain the experiments that you have in your setting and to show the effect of prior authorizations for specific common drugs. This gets the readers on board, and will provide a nice segue into the full regression model.*

Thank you for this terrific suggestion, it indeed is a really nice prologue to our regression approach. We have implemented it exactly as you suggested, both for the effect on the focal drug and the effect on substitutes. This is now in Figures 5a and 5b. We describe the exercise in Section 3.2:

> *As stated above, our approach is equivalent to an average over many within-drug comparisons. Before running regressions, we demonstrate some of these comparisons to show that the effects of prior authorization are apparent even when looking at the data in its rawest form. In Figure 5, for a subsample of drugs, we plot the share of beneficiaries who filled the drug at least once in the year (after residualizing market fixed effects), comparing those assigned to a plan that restricted the drug versus those who faced unrestricted coverage of the drug. Unsurprisingly, for each drug, when a beneficiary faces a prior authorization restriction, they are less likely to ever fill the drug. The regression estimates we discuss below can be viewed as weighted averages over many of these drug-specific comparisons.*

and Section 3.4:

> *We start by providing evidence of the existence of substitution effects and proceed to estimate the extent of those effects. To demonstrate the existence of substitution, Figure 5b presents results similar to those presented in Figure 5a but focusing on substitution instead of use of the focal drug. Specifically, for a subset of drugs, we compare plans that restrict the drug to those that don't. However, in this figure we plot the use of other drugs in the same therapeutic class as the listed drug for beneficiaries randomly assigned to plans that place prior authorization restrictions on the listed drug versus beneficiaries randomly assigned to plans that do not restrict the listed drug (residualizing market fixed effects). We see that beneficiaries who face restrictions on the listed drug (who, as shown in Figure 5a are less likely to consume the listed drug) are more likely to*

> *consume other drugs in the same class as the listed drug, relative to beneficiaries who do not face restrictions on the listed drug. These results clearly indicate some degree of substitution due to prior authorization restrictions.*

(E.2f) *All estimates should have standard errors. This includes estimates of loss in CS and estimates of net spending reductions on drugs.*

This is now true for all tables in the paper.

(E.2g) *I wasn't totally clear how your usage variable was defined. Was it any usage in a given year or did the usage measure capture for how many days the drug was prescribed that year? Both usage metrics might be interesting. Please discuss in the paper which usage variable(s) you are using and why.*

Our utilization variable in the original manuscript was a binary indicator for whether the drug was ever used in the year. As mentioned before, our main regression table now also includes specifications where the outcome is the yearly days supply and yearly spending for the drug as well. These outcomes were in the prior submission but in an appendix table. See descriptive text below on page 17:

> *We focus on three different measures of utilization: A binary indicator for whether the beneficiary filled the drug at least once in the year (multiplied by 100 to reflect percentage point changes), a count of the total days supply of the drug filled in the year, and total allowed spending on the drug (net of rebates).*

(E.2h) *The controls for prior authorizations or exclusions of substitute drugs did not seem to make any difference. If so, why not leave them out? This simplifies notation, especially for the first stage.*

While you are correct that they did not matter in the primary regressions used in the original manuscript, they are *essential* when estimating effects on substitutes in the specification added as a result of comment E.2, since the outcome measure includes the use of drugs which may themselves be restricted. If we exclude the control in those specifications, we get wrong-signed estimates (i.e., we would erroneously conclude that prior authorization on drug A *reduces* the use of drug B/C/D/E). We have therefore left them in our primary specification in order to have a single consistent specification throughout.

(E.2i) *Please use IV or ITT results throughout. Given how precise the first stage is, I didn't buy your argument that for some specifications you reverted to ITT because IV made standard errors less precise. I'm also fine if you report ITT estimates throughout and simply say the first stage is 0.9, so the IV estimates are 10% larger than the ITTs. The first stage is so strong that the first stage does not need a lot of space or attention in the paper. I'd think a few sentences maximum, not an entire section.*

We have shifted all estimates to being IV estimates when relevant. This change primarily affected the estimates of health effects, which were ITT estimates in the original manuscript. We have also folded what was previously Section 3.2 ("First Stage") into the prior section on research design and reduced the discussion of Table 4 substantially.

(E.3) *Willingness-to-Pay estimates. For each drug, you just have two points at the demand curve, the usage fraction at the regular price and the usage fraction with a price of zero. Is there any way you can do these estimates using the same linear model as in (2)? Why or why not? Rather than imposing the same semi-elasticity for all drugs, could you allow for a different slope by therapeutic class? What is the role of functional form assumptions here?*

Thank you for this suggestion. We have updated the WTP estimates accordingly. Because individual fixed effects and drug-market-time fixed effects are necessary in order to identify the effects of joining the LIS (where our variation in price comes from), we do this via regression rather than just taking the raw observed spending in and out of the LIS. Specifically, we replicate the specification of Gross et al. (2022) and run two *linear* regressions, one where the outcome is the out-of-pocket price paid for prescriptions and another where the outcome is utilization. This provides us with an estimate of the change in price when joining the LIS and the change in utilization. We combine these estimates with the average price and utilization prior to joining the LIS to do exactly as you suggest and produce two points on the demand curve. We then extrapolate the demand curve to the y-axis and use this demand curve to estimate average WTP of the marginal beneficiaries under the various assumptions we used in the prior submission (perfect screening, random screening). These results are reported in Tables 9 and 10. We find these results clearer and more transparent and thank you for the suggestion.

We also include several robustness checks, including the non-linear (semi-elasticity) specification from the prior submission and the specification you suggest where we allow the slope of the demand curve to vary across therapeutic classes.

(E.4) *Health effects estimates. I appreciate the attempt at the case study of anti-coagulants, but as you write, we don't learn much from it. The almost full substitution towards the generic means that even if you had statistical power, it would be a case where we would not expect health effects. I think you should cut this case study. The overall health estimates are helpful, even if*

they only serve as a caveat on a potentially important unmeasured effect. I understand why you use yet a different specification (and here you indeed need to restrict the sample to those whose plan ceased to be a benchmark) and why you move to quintiles. However, is there no way to stay closer to the original specification and make it feel less ad hoc? E.g., the quintiles absorb a lot of useful variation. Why quintiles? Why only look at the highest one? How about a specification where the explanatory variable is the number of prior drugs used that require an authorization in the current plan minus the number of prior drugs used that require an authorization in the average plan. This way the explanatory variable is random with mean zero by construction, but you retain more power, it feels less arbitrary, and the magnitudes are more easily interpretable.

We understand your objections to the inclusion of the oral anticoagulants analysis. In early presentations of this project, we received many requests to estimate effects on health in different ways (comments similar to R2.3), and this was one of the ways we thought was plausible. We agree that the case study is not especially revealing, though. As suggested, we have cut it.

We used the quintile specification in what was previously Section 6.2.2 to follow the analysis from our prior paper that we borrow the design from. However, your suggestion was very reasonable and we agree that it is a more straightforward approach that we had not thought about. We have implemented your precise suggestion and included it as our primary estimate in what is now Table 11. This specification, like the prior one, finds that enrollment in plans that restrict more previously-used drugs reduces spending on drugs. In contrast, we find that it *decreases* mortality and non-drug spending, though, as before, our standard errors are large. (Note that, as requested in comment E.2i, we have converted this to an IV estimate so that we are consistent with other exercises in the paper).

(E.5) *I would like you to be clearer to about how your estimates are aggregated if there are heterogeneous effects. Which one are weighted across drug (classes) by precision of the estimates, which ones are weighted by spending, etc.? Conceivably, you could provide both if you do the estimates separately for each drug class (see comment 2d), you can then choose how to aggregate them (by spending, by precision, or perhaps by some other relevant metric). The weight expression at the bottom of page 26 did not make sense to me. If the variances were taken across individuals, the resulting weight should not be individual specific. Or were the variance calculated across something else. Please specify this.*

Thanks for this comment. We think the confusion is about notation. We have changed the notation to (hopefully!) make the weights clearer. They are unique at the drug-market level, where a market is a region-year pair, which depends on i and t, thus making the weights beneficiary-drug-year-specific. It is indeed true that all of our headline reduced-form estimates represent across-drug-market weighted average treatment effects, with drug-market-level weights.

We now discuss the weights more thoroughly in Section 3.1. It previously appeared much later, we

think this placement is more appropriate. We say:

> *The implicit weights in the OLS estimation of our effects reflect how precisely each drug-specific difference is estimated (Gibbons et al. 2019). Specifically, the weights are equal to* $w_{idt} = p_{dm(it)}(1 - p_{dm(it)})$*, where* p_{dm} *is the probability that a beneficiary in market* m *will be assigned to a plan that restricts drug* d*. For instance, if a drug never faces prior authorization, we can never estimate the effect on its use; if 50% of beneficiaries face prior authorization, we can estimate drug-specific effects more precisely. The weighted average treatment effect, with these weights, is estimated more precisely than the unweighted average treatment effect. All specifications we present in this paper use these weights, either implicitly or explicitly.*

In the process of this revision, we played around with the idea of producing unweighted estimates (i.e., putting weights within our regression approach to undo the implicit regression weights). This produced wild and extremely noisy estimates. This occurs since there are many drugs where we have virtually no variation, e.g. they are *almost* never under prior authorization, and thus the associated (implicit) standard errors for the effect of prior authorization on their use are very large. We also considered weighting by, e.g., spending or baseline use, but it felt awkward to use weights that were also left-hand-side outcomes. Our hope is that now that we use a single consistent specification for all main results that there is less confusion about how weights may or may not differ across specifications.

(E.6) *The revealed preference notation seems needlessly complicated for what is essentially a demand curve and a change in consumer surplus. Is there any way to present this in a way it does not look more complicated than it actually is?*

We used this notation in order to be consistent with the notation used for the model in Section 1.2. We have now edited it to severely reduce notation, except where (we felt) it was absolutely necessary. We hope it is clearer now.

(E.7) *Please spend some time on making equations, tables, and figures more user friendly. Please write out what variables are rather than use abbreviations (to the extent possible), make clear in the tables itself what numbers in different rows mean, add white space when it helps with legibility, and have self-contained notes, etc.*

We have revised our exhibits substantially to do this. More importantly in our view, we have replaced all the regressor labels in our tables to reflect text descriptions of the regressors rather than generic variable names.

(E.8) *Guidance on comments by reviewers*

a. *(R1) R1's comments can in my view all be addressed by adjusting the write-up. I'm looking for relatively minor adjustments in the write-up: changing some of the existing language, a few well-chosen additional sentences, but not (lengthy) additional paragraphs on each of the reviewer's comments.*

b. *(R2.2) R2 comment 2. This is a great point. You can show this distinction in some of the heterogeneity analysis, perhaps even in the graphical evidence I suggest in my comment 2a or in the price responsiveness estimates, see comment 8f below.*

c. *(R2.7) R2 comment 7. I would like to keep the revealed preference approach in the main paper.*

d. *(R3.1) R3 welfare analysis. This reviewer is technically correct if the goal is a full social welfare analysis. However, a full social welfare analysis of a wholesale policy of banning prior authorizations (or requiring them for all drugs) is close to impossible anyways due to likely general-equilibrium effects. I would suggest you explicitly limit your ambitious and describe the exercise in more modest terms. I would focus, as you do, on the partial equilibrium effects from a single plan changing a single prior authorization taking into account effects on those directly involved, namely patients and doctors. Please retain as caveats differences between effects that you include in your limited welfare analysis and additional effects that would need to be taken into account into a full-scale global social welfare analysis.*

e. *The second editor also remarked on the welfare analysis, finding it "a bit sketchy." I myself would call it heroic, but the underlying idea is the same. As you recognize, it is extremely hard to do a credible full-blown welfare analysis in your setting. However, I still see value in what you do in Section 6: it discusses what is needed for a welfare analysis, provides the reader with a sense of orders of magnitude, and it gives readers a sense that the loss in CS is probably smaller than the savings in spending net of admin cost (subject to a bunch of caveats). I strongly encourage you to present the welfare analysis in the spirit that I describe here, not just in Section 6, but also earlier in the paper and in the conclusion. This would also mean being careful about (or avoiding altogether) the use of terms such as "inefficient" or "welfare". E.g., instead of inefficient, you could say that admin cost and the lost consumer surplus are larger than the reduction in spending.*

f. *The second editor pointed out an additional caveat in the welfare analysis. If some people react to copays more than others, the lost consumer surplus is larger than you calculate. See Alex Rees-Jones & Dmitry Taubinsky, "Attention Variation and Welfare: Theory and Evidence from a Tax Salience Experiment" (REStud 2018). This may be a worthwhile caveat to point out. In addition, you could get at different elasticities like R2 comment 2 suggested, namely by "seriousness", and do the CS calculation separately for these two*

classes of drugs. This would be at least one step towards recognizing some heterogeneity.

g. *Finally, the second editor pointed out that the sample on which you estimate the price semi elasticity is different (likely richer, more used to looking at copays) from the sample on which you estimate the effects of prior authorizations. This difference is worth mentioning as a caveat.*

h. *(R3.3a) R3, comment 3a – See above, I strongly suggest removing the nested logit altogether (which fits with the spirit of this reviewer's comment).*

i. *(R3.3b) R3, comment 3b – This is a valid point, but the WTP is illustrative, so rather than making it more complicated, I would acknowledge this limitation in the write up.*

j. *(R3.2) R3 counterfactual – I think this can be addressed in write-up. See my comments 8d and 8e above.*

k. *(R4.4) R4, comment 4, on WTP estimates. I agree with this comment. My comment #3 can be seen as a reinforcing this point.*

Thank you for your guidance on responding to the reviewers' comments. We have responded to them in the respective sections of this document. We have embedded the relevant hyperlinks above so you can easily navigate to the responses.

On comment (e): Your and the second editor's comment about the language used to describe our welfare analysis makes sense. We have edited that section substantially to tone down language that was overly definitive about interpreting our results strongly as welfare. Here is a snippet from the intro to Section 5 where we motivate the welfare analysis section while also providing appropriate caveats regarding its usefulness:

> *For a variety of reasons, it is difficult to estimate this lost patient surplus. The two primary approaches for estimating patient surplus involve (1) inferring patient valuation from patient choices (revealed preference) and (2) inferring valuation from estimates of health effects combined with the value of a statistical life-year. In our setting, (1) is difficult due to the fact that LIS beneficiaries do not face non-zero prices for filling prescriptions, limiting our ability to assess patient willingness-to-pay for drugs. Further, when patients do face positive prices for healthcare, there is substantial evidence that demand for drugs and services may not reflect clinical or private value (Baicker et al., 2015; Brot-Goldberg et al., 2017; Chandra et al., 2021). Similarly, (2) is difficult due to the fact that the typical beneficiary-level measure of health that is available is mortality, and a marginal reduction in drug utilization, especially for the categories of drugs which are frequently restricted, may take years to have a meaningful effect on mortality rates.*
>
> *Despite these difficulties, we argue that it is a useful exercise to attempt to estimate*

> *lost patient surplus using variations on these methods in order to (1) provide guidance for how one might evaluate welfare consequences of this type of "rationing via bureaucracy" and (2) provide a kind of benchmark of what might be a reasonable guess for the reduction in patient surplus due to prior authorization. We do so in this section.*

See also our response to R1.2 who objects to our use of the phrase "moral hazard" which has similar welfare connotations. We have toned down our references to moral hazard as well.

On comment (f) from the second editor: We have included a version of this analysis where we estimate class-specific price-sensitivity parameters and aggregated up from that level, in the spirit of the Rees-Jones-Taubinsky reference. We have also included language in this section to reference this point on page 30:

> *As Taubinsky and Rees-Jones (2018) highlight, aggregating effects while ignoring heterogeneity in demand responses may underestimate total beneficiary value. Therefore, in Column (2), we present a version where we estimate therapeutic-class-level responses to the LIS transition and use them to construct class-specific demand curves, then aggregate up.*

Reviewer 1

(R1.S) *My impression is that insurers frequently place prior authorization restrictions on prescription drugs, particularly in insurance systems with little or no patient cost sharing. Despite being a simple mechanism at first glance, prior authorization restrictions likely play a complex role in insurance design. They involve multiple parties (i.e., patients, providers, insurers, and PBMs) and various potential costs and benefits and are potentially an important component of many complex interactions and negotiations. Admittedly, I worry that the paper may sidestep important parts of the complexity. However, I recognize it likely does so to gain tractability. Moreover, I think the paper does an impressive job of providing a framework that is both useful for its particular setting and will be useful for other studies of prior authorization restrictions.*
The paper is ambitious in its scope and attempts to assess various costs and benefits of prior authorization restrictions. I think the paper does a good job of considering first-order issues with prior authorization restrictions, and I generally found the empirics convincing. But given the ambitions of the project, a few big picture issues seem relevant that I think the paper could guide readers on a bit more. I mention them and a few other specific thoughts below.

Thank you for this generous summary. We agree that there is a lot of complexity we do not cover in this paper, and we appreciate that you understand the difficulty of tackling it! We hope this revision made the paper better in your eyes. We note that the guidance from the Editor (E.8) was to respond to your high-level suggestions (at least, within the manuscript itself) briefly when appropriate.

(R1.1) *Determination of prior authorization restrictions. I think it's fine that the paper's formal frameworks exclude insurer decisions (again, I recognize how complex of a setting the authors are attempting to shed light on here!), but a more explicit understanding of how/why insurers make decisions about prior authorization restrictions would have been helpful for me. I imagine inefficient prescriptions raise plan costs without raising plan value, which would drive up premiums and cause plans to lose market share. Is this right? Otherwise, insurers would not care about costs, right? Any guidance the authors could provide here would be helpful.*

a. *Relatedly, I also think prior authorization restrictions could function as a way to get high-cost enrollees off a specific plan (which is related to moral hazard comment below). I know the authors show compliance with plan assignment is high, but I think we might expect the relevant non-compliers for a specific drug restriction to be few but important.*
b. *Relatedly, I do not understand why there is much variation in prior authorization restrictions across plans. With the Lipitor discussion, for example, it really felt like I was missing something with the decision about how insurers designed prior authorization restrictions since there is apparently not a dominant design.*

Thank you for this helpful comment. We now address part of this in a footnote:

> *Note that screening need not be the only motivation for prior authorization. Insurers may also impose restrictions to discourage users of the restricted drug from enrolling in their plan (Geruso et al. 2019). They may impose restrictions on rival drugs as a reward for rebate payments (Brot-Goldberg et al. 2022, Ho and Lee 2023).*

To your point b: While there is no full theoretical characterization of when this might occur, it is plausible that the rebate concession motive might generate separating equilibria where insurers pursue different strategies (e.g. insurer 1 caters to manufacturer A while insurer 2 caters to manufacturer B), leading to heterogeneity in formularies. Another possibility is that some plans are simply better than others at efficient formulary design. Further, some of the variation may come from the 'path to equilibrium,' as different insurers observe the restrictions imposed by their competitors and respond. Even if all insurers end up at the same place, it may take a few rounds of the game to get there, generating within-drug-market-year variation in restrictions across plans along the way. We have added a few sentences to the end of Section 2.4 explaining these ideas:

> *Importantly for our identification strategy, prior authorization varies significantly across plans for a given drug. For each drug, in each region and year, we compute the share of offered benchmark plans that restricted that drug. Figure 4 displays the distribution of this share across drug-region-years, omitting cases where the share is 0 or 1, which comprise 74.2% and 2.6% of drug-region-year tuples, respectively. We observe full support across the* $[0, 1]$ *interval. The reasons for this variation are unclear, but we provide two potential explanations. First, it is plausible that the rebate concession motive (firms offer larger rebates in order to avoid restrictions) might generate separating equilibria where insurers pursue different strategies (e.g. insurer 1 caters to manufacturer A while insurer 2 caters to manufacturer B), leading to heterogeneity in formularies. Second, some of the variation may come from the 'path to equilibrium,' as different insurers observe the restrictions imposed by their competitors and respond. Even if all insurers end up at the same place, it may take a few rounds of the game to get there, generating within-drug-market-year variation in restrictions across plans along the way.*

We would not assert that plan design is necessarily optimal here! Fortunately, our analysis does not require this.

(R1.2) *Moral hazard. I felt a little uneasy about the use of the term "moral hazard" throughout the paper. In the end, I think I maybe agree with the authors that it is fine to think of the purpose of prior authorization restrictions as being to reduce moral hazard. But it's strange for me to*

think about insurers' approach to ensuring that patients truly value the drugs they are receiving being one that taxes providers (as in prior authorization restrictions) unless providers are perfect agents for their patients, but the paper seems agnostic about the degree to which providers act as perfect agents for their patients. I know the appendix discusses the role of agency in more detail, but perhaps a few additional sentences could go in the main text. I wonder if prior authorization restrictions have heightened potential to be inefficient (perhaps more type II errors?) if providers are imperfect agents. Also, as noted in the previous comment, it seems like insurers would try to use prior authorization restrictions to influence the types of enrollees who enroll in their plans.

Rereading the paper, we understand this sentiment. We think there were a few offending sentences that we have now changed:

> ***OLD (p.1):*** *We conceptualize prior authorization as a tool for insurers to fight moral hazard problems, where generous insurance coverage may incentivize the use of low-value care (Pauly 1968).*
>
> ***NEW (p.1):*** *We conceptualize prior authorization as a tool for insurers to fight moral hazard problems and reduce the use of low-value care. Prior authorization forms allow providers to directly communicate information to insurers about the patient's suitability for the drug, helping resolve a key information asymmetry and allowing insurers to target coverage denials to low-value use.*

> ***OLD (p.6):*** *Since patients do not internalize social costs, under this choice utility function...*
>
> ***NEW (p.6):*** *Since neither patients nor providers internalize social costs, under this choice utility function...*

> ***OLD (p.6):*** *...the classic case of moral hazard (Pauly 1968).*
>
> ***NEW:*** *[no longer in the paper]*

> ***OLD (p.29):*** *Therefore, in these two cases, this exercise suggests that that there is enough moral hazard among these beneficiaries such that the reductions in utilization are Kaldor-Hicks efficient...*
>
> ***NEW:*** *[no longer in the paper]*

> ***OLD (p.29):*** *Ultimately, we interpret the results from this revealed preference approach as suggestive evidence that the lost consumer surplus may be sufficiently low relative to the cost of care to indicate substantial moral hazard in this setting. This moral hazard motivates prior authorization restrictions as a potentially efficiency-enhancing rationing device.*
>
> ***NEW (p.32):*** *Explicitly interpreting these estimated quantities as reflecting the welfare consequences of prior authorization requires strong assumptions. This is reflected in the wide range of estimates we obtain under different approaches. Ultimately, we interpret these results as suggestive evidence that the lost consumer surplus due to forgone drugs is of a similar order of magnitude relative to the cost of procurement. This motivates prior authorization restrictions as a potentially efficiency-enhancing rationing device.*

The edits (particularly the first two) now emphasize that prior authorization need not only solve patient-side moral hazard problems, it could also solve provider-side moral hazard (i.e., providers have incentives to overprescribe), the latter being a more natural justification for a provider-side remedy. (For this reason, we took the Pauly cite out of pages 1 and 6, since it only focuses on patient-side moral hazard)

Your broader comments about imperfect provider agency are very interesting, and we think it is a fruitful line of thought. But we also agree with your earlier comment that the paper is already a bit sprawling, and we didn't feel like we had a great way to empirically analyze imperfect provider agency, nor talk about when it does or does not justify prior authorization as a policy.

(R1.3) *Prior authorization and prices. My understanding is that prior authorization restrictions are crucial in negotiating drug prices and rebates. Insurers/PBMs may exchange the removal of prior authorization restrictions for lower prices from drug manufacturers. Because of data limitations, it may not be possible to empirically assess the role of prior authorization restrictions in determining net prices, but I was wondering if the paper could do more to guide our thinking here a bit more.*

a. Relatedly, the paper largely sidesteps issues around PBMs vs. insurers. On the one hand, it felt like a wise decision not to get bogged down in part of the process that is so opaque, but it also left me wondering if maybe PBMs weren't as relevant as I thought. Again, any guidance would be useful.

We agree that PBMs are very important in determining prices and formularies, but since we are focusing on patient/provider responses rather than insurer strategy, we didn't think it was important to discuss them at length. We don't assess effects on net prices since our design isn't suitable for it without additional structure on supplier behavior. To assess that in a reduced-form way, we would

need (quasi-)random assignment of drugs to prior authorization, which we do not have. Instead, we observe equilibrium prices and prior authorization restrictions. We hope that future researchers will be able to say something about this.

(R1.4) *Heterogeneity. The heterogeneity analysis is useful. I wonder if the willingness to fill out prior authorization requests varies systematically across providers. It might be difficult to interpret if doctors who do not fill out prior authorization requests (if there are such doctors) have different patients, but it is also difficult to imagine that it would be efficient for some doctors to blanketly refuse to fill out prior authorization requests. Perhaps considering how the effects of prior authorization restrictions vary across doctors is another way to test for inefficient effects of prior authorization restrictions.*

At the start of this project, we were also very interested in the heterogeneous effects of prior authorization across different providers. The dilemma we face is the following: Our variation is across beneficiaries, so to do this we would have to assign beneficiaries to providers. A common way is to assign based on the modal doctor visited. But this is typically used to assign patients to primary care providers, who we think are unlikely to be prescribing some of the most heavily-restricted drugs (e.g. we think cancer drugs are likely being prescribed by oncologists), so we would need to take a stand on the likely potential prescriber for any given drug, which is especially tricky in the cases where the patient does *not* receive a drug. This is a natural weakness of studying drugs, for which the prescribing provider does not receive direct reimbursement. Given the current paper scope, we thought this was too complex an exercise to merit inclusion.

(R1.5) *Comparison to prior authorization requirements more broadly. In practice, the specifics of the prior authorization restrictions likely matter, in terms of both prior authorization restrictions' costs and benefits. A system that is highly onerous or that nearly always rejects will have different effects than one that is trivial. I appreciate the authors providing prior authorization request forms in the appendix. My impression is that there is not a lot of variation in most prior authorization request systems in the United States and that prior authorization in the insurance system studied here is standard. If that is correct, perhaps the authors could mention. Or perhaps the authors could include a few sentences about how the prior authorization details here compare to prior authorization in other insurance systems.*

Thanks for this comment. We have now included this in a footnote in Section 1.1:

> *Both the forms and the sequence of events required for authorization for drugs, or for other services, are broadly similar for insurers in other U.S. insurance market segments.*

As an example, here is a request form for imaging: https://www.huskyhealthct.org/providers/provider_postings/provider_forms/Advanced_Imaging_PA_Request_Form.pdf

(R1.6-7) *6) Health impacts. I appreciate that the authors are careful when assessing health impacts to recognize two inherent limitations of doing so: 1) They must focus on case studies, and 2) They cannot rule out meaningful impacts. Despite these limitations, I appreciate the authors attempting to consider health impacts because they are potentially a relevant cost of prior authorization restrictions and because the study's setting allows for highlighting the difficulties of conducting meaningful analysis of the impacts of prior authorization restrictions on health.*
7) Conclusion. I think the conclusion is great. Throughout, there were several issues that I thought were perhaps beyond the scope of the paper but were also maybe too relevant to ignore. I thought the conclusion did an excellent job of connecting the results of the paper to broader issues.

Thank you for these kind comments!

Reviewer 2

(R2.1) *The model in section 1.2 is helpful for fixing ideas about the use of prior authorization (and supply-side rationing mechanisms more generally) within a normative framework. I found this paragraph particularly helpful:*

> *The ideal drug to restrict, from this perspective, is an expensive, niche branded drug, especially one that is a new entrant within an established therapeutic class. The worst are those like generic aspirin: Drugs which can be cheaply procured, have high incremental patient value (since the next alternative is likely to be nothing), and substantial numbers of inframarginal users. One caveat applies: If the incremental net social value of a drug is too small (e.g. expensive branded drugs with cheap bioequivalent generic substitutes, where there is little justification to purchase the branded option), prior authorization will be too weak a tool to use to improve social welfare since it may still permit uses of the drug, essentially all of which are inefficient. In that case, a policymaker should want to exclude the drug from coverage outright.*

My question is how to connect the ideas from your model with one pattern that is borne out in the data. On page 12 you say:

> *The least-restricted drugs are branded drugs with generic bioequivalents. This is because, as suggested in Section 1.2.1, prior authorization is too weak a restriction for such drugs. The average drug in this category is, instead, excluded in 57.2% of plan-years.*

Your model suggests that drugs of this type deliver low value in the market given high list prices and the availability of (essentially) perfect substitutes at lower prices. It makes sense then that a majority of plan-years choose to exclude these drugs from the formulary all together. But among the 43% of remaining plan-years, why wouldn't insurers place a prior authorization restriction? The conclusion that exclusion is preferrable to prior authorization makes sense to me, but then surely this would also imply that prior authorization restriction is preferrable to no restriction? Why are these plans leaving these drugs unrestricted? I am wondering about this because I worry that there are many plans (maybe up to 43% of plan-years) where there are very few restrictions on formularies of any type, and the plans attempt to either (i) control utilization in other ways, or (ii) charge higher premiums and position themselves as "high quality" plans, or (iii) choose not to exclude or restrict these branded drugs as part of a negotiated agreement with manufacturers over rebates. Could you provide some interpretation of this pattern and discuss how it might interact with your results?

Thank you for this comment. We agree that this insurer strategy seems suboptimal on its face. We will respond here, but our responses are mostly speculative so we did not include them in the revised manuscript (though we are happy to do so if you/the editor feel strongly about it).

First, it is worth remembering that the plans we study are serving both the LIS population (which we study) and the non-LIS population (which we do not). So, for example, the plan might cover these drugs but require substantial copays for them, copays that don't matter for the LIS population for whom copays are fully subsidized. Your explanation (i) seems unlikely, since there *aren't* other non-price barriers insurers can impose in this setting (as demonstrated by our estimates' robustness to including plan FEs as well as richer plan FEs interacted with various drug characteristics). (ii) and (iii) are plausible explanations that would have no bearing on our results. For (ii), recall that our instrument only assigns beneficiaries to zero-premium plans, so we need not worry about how premiums affect behavior. For (iii), while prior authorization might affect rebates, LIS beneficiaries never face prices for covered drugs anyway, so this would not induce any relevant unobserved variation.

Both you and R1.1 requested a little bit more on how plans set formularies, so we have added some text in Section 1.1, though we have tried to keep it minimal since we do not study supply side behavior:

> *Note that screening need not be the only motivation for prior authorization. Insurers may also impose restrictions to discourage users of the restricted drug from enrolling in their plan (Geruso et al. 2019). They may impose restrictions on rival drugs as a reward for rebate payments (Brot-Goldberg et al. 2022, Ho and Lee 2023).*

and at the end of Section 2.4:

> *We cannot definitively explain the reasons for the remaining variation, but we provide two potential explanations. First, it is plausible that the rebate concession motive (wherein manufacturers offer larger rebates in order to avoid restrictions) might generate separating equilibria where insurers pursue different strategies (e.g. one insurer caters to manufacturer A while another insurer caters to manufacturer B), leading to heterogeneity in formularies. Second, some of the variation may come from the 'path to equilibrium,' as different insurers observe the restrictions imposed by their competitors and respond. Even if we might expect a symmetric equilibrium, it may take time for insurers to reach this point, generating within-drug-market-year variation in restrictions across plans along the way.*

One thing we want to emphasize in our response is that even though we argue that prior authorization, as applied in our setting, is not necessarily welfare-decreasing, we are *not* asserting that insurers are necessarily setting formularies optimally.

(R2.2) *I'm a bit surprised by your estimates that so many people substitute to "no drug" when faced with a prior authorization restriction. Your explanation is that the restriction places a burden on patients, who decide to just not fill the drug instead of asking their doctor to provide the authorization or else a different prescription. In my experience, the prior authorization process can happen without me doing anything: the PBM contacts the prescribing physician to ask for authorization paperwork, which then gets completed, at which point I can fill my prescription. I understand this may not be representative of the experience of Medicare LIS enrollees, but it also seems possible that it is.*
Would you be able to provide a breakdown of which types of drugs are driving the strong "extensive margin" effects of substitution to "no drug"? Do you see stronger "no drug" substitution within "less important" therapeutic classes (topicals?), and stronger "alternative drug" substitution within classes that seem critical for maintaining good health?

These are useful thoughts. However, we think your experience might not be representative during our sample period, for two reasons. First, part of the speed of response is that providers now may have electronic prior authorization request systems integrated into their IT system. These did not exist during our sample period. We now mention this in a footnote in Section 1.1:

> *In our sample period, forms were primarily sent via fax for legal compliance reasons. In later years, some providers were able to obtain electronic prior authorization assistance integrated into their IT apparatus.*

Second, these systems are now more common in more profitable health systems, which are more likely to serve academics like us and less likely to serve low-income patients like those in our sample.

We did estimate heterogeneous effects on diversion by class (see our response to E.2 for more details on this exercise). Unfortunately, because class-level utilization is quite noisy, our diversion results (which take the ratio of two class-level estimates) are even noisier and probably are distributed Cauchy (i.e. with undefined mean and variance). We did not include them in the paper because we felt they were so uninformative.

(R2.3) *Regarding the effects on health outcomes, my prior is that any strong health effects would be operating through patients who substituted to "no drug" as opposed to an alternative drug. This would be consistent with the fact that in Section 6.2.1 you find no measurable health effects among anticoagulants, since you report that there is essentially no extensive margin substitution within this class.*
If you are able to isolate which types of drugs saw the largest extensive margin substitution (from point 2), would you be able to look for health effects in those places specifically?

As per our response to E.4, we no longer include the oral anticoagulants case study in the paper. Your suggestion makes sense. The development of that case study came out of a long process by which we interviewed doctors and asked them what drug classes might plausibly (1) have frequent prior authorization burdens (i.e., not drugs like aspirin); (2) have distinct health consequences that could plausibly be measured in a short time span (i.e., not cancer drugs); and (3) have enough baseline users that analysis could plausibly be powered (i.e., not niche specialty drugs, though our case study was not powered in the end, either). Oral anticoagulants were the only class that appeared to satisfy all three. Further, given our overall difficulties getting enough power to estimate aggregate health effects, we felt like it would be even more unlikely that we could achieve sufficient power to say whether health effects were larger for some drugs versus others.

(R2.4) *The authors discuss a number of analyses that turn up inconclusive results. For example, the three paragraphs at the end of section 3 on page 18, as well as the analyses of health effects. These analyses are extremely logical to have explored, but given the very noisy estimates, I wasn't sure how much could be taken away from them. It may make the paper more streamlined to include a briefer mention of these analyses in the main text and move the bulk of the text to an appendix.*

Thanks for this suggestion. As requested by the editor, we have removed the anticoagulant analysis from the paper, though he requested we keep the aggregate health effect estimates. We have also reduced the discussion at the end of Section 3.

(R2.5) *On the top of page 12, the authors describe the comparison between plans that enrollees are randomly assigned to versus plans that enrollees actually enroll in: "Plans that beneficiaries actually enroll in generally look similar, in aggregate, to plans that they are assigned to." I was a bit confused about why the plans people actually enrolled in weren't "better" in some way that the plans to which they were randomly assigned to? I would have thought that the non-compliers (those who do actively select a plan) are opting for plans that seem better on observable dimensions. Why is that not the case?*

Sorry for the confusion, this was the result of sloppy writing. The reason the two groups of plans are similar is precisely *because* compliance is so high, and we do not restrict to only the non-compliers when making the comparison, so naturally these two groups of plans are similar because they are mostly the same. We have removed the sentence so as not to confuse future readers.

Additionally, we should note that most beneficiaries in this sample, when making active choices, choose other zero-premium benchmark plans (this is not in the manuscript, but we document it in our prior paper, "The Behavioral Foundations of Default Effects"), so the chosen plans are more likely to be horizontally differentiated than vertically differentiated. This shouldn't matter for our estimates, however, since our estimates are LATEs only for the compliers.

(R2.6) *You mention that prior authorization is in some instances used for safety reasons. Would there be a way to drop all such drugs from your analysis, so that you focused only on drugs for which the motivation for prior authorization is to reduce low-value utilization?*

We do this in Section 3.3 by separating the "scheduled" and "unscheduled" drugs. We think this is the best proxy we have for which drugs are restricted for safety reasons, with the "unscheduled" being those who are not likely to face prior authorization for safety reasons. As shown in Figure 8, the estimates for the "unscheduled" are nearly the same as the overall estimates, suggesting that the main results are not driven by the scheduled drugs that are restricted for safety reasons (and actually have smaller effects of prior authorization).

(R2.7) *The "revealed preference approach" in section 6.1 is somewhat hard to interpret. The assumption that drug consumption would otherwise transition smoothly upon qualification for the LIS program seems suspect. Beyond that, there are many other, more difficult to interpret, assumptions baked into the model in order to make it tractable to estimate. Ultimately, I felt that the whole analysis didn't shift my prior much one way or another with respect to the value enrollees derive from prior-authorization-restricted prescription drugs. At the bottom of page 29, you discuss an alternative approach in which one considers the perspective of the physician deciding whether to prescribe a restricted drug (and pay the associated administrative cost) versus a non-restricted drug. This analysis seems in many ways more straightforward. I would encourage you to consider placing this analysis in the main text, and instead placing the patient-side revealed preference analysis in the Appendix.*

Thanks for these thoughts. Our main reason for emphasizing the beneficiary revealed preference results over the provider revealed preference results was that the provider results always suggested even lower WTP than the beneficiary results, so the beneficiary results represented the kind of upper-bound for our WTP estimates. We're sympathetic to your preference for the provider revealed preference results, but the editor has requested that we tighten up this section so we opted to continue to emphasize the beneficiary-side results (see E.8), though we do include a discussion of the provider results at the end of Section 5.1. As per the editor's request, we have significantly edited this section to make much clearer that this exercise requires substantial assumptions.

(R2.8) *Despite the disappointing results on health effects"... My take was actually that the lack of strong evidence of negative health effects was a very positive outcome!*

Fair point! We have tried to be very clear in this section about what we can and cannot say. We can say that there is no clear evidence of health effects of prior authorization. But we can't say that there is evidence that there are necessarily no health effects of prior authorization. We discuss this in

the revised version of the manuscript, emphasizing the wide confidence intervals on our estimates that include both positive and negative effects on health.

(R2.9) *The conclusion is a bit long and repetitive with much of the earlier text.*

We have shortened the conclusion slightly, focusing on the spots where we summarized our own analysis rather than where we talked about broader implications.

Reviewer 3

(R3.1) *Welfare Consequences: The paper makes an overwhelming case that prior authorization reduces health care spending, net of administrative costs, relative to a world with no prior authorization and the same plan formularies. But I do not think the welfare calculation considers the right object on the cost side or, possibly, on the benefit side.*

(a) To assess overall effciency or welfare impacts, the change in "cost" is not the change in drug spending, which is simply a transfer from Medicare/insurers to pharmaceutical companies. One could argue about exactly how the cost side should be formulated, but it should include marginal resource cost of producing the drug, plus the social cost of funds from distortionary taxation to raise revenue to fund the transfer (assuming Medicare foots the bill in the end). The paper's analysis assigns zero welfare weight to pharmaceutical company revenue; this alternative version would assign equal weight to all economic agents. If the authors believe their version is the right one, that at least requires justification. Either way, the "cost of funds" should also be part of the welfare calculation (using calibrated values from the literature).

Thank you for this comment. We had previously discussed this a bit in a footnote in the theory section. We now discuss it in more detail in the main text, quoted below. Upon a re-read, we realized that some of our discussion of "costs" was sloppy and so we have tightened it up. Practically, we agree with you that a full view of social welfare would require us to take manufacturer profits and the cost of funds seriously. In practice, we don't have estimates of manufacturer marginal costs so we cannot estimate manufacturer profits. As we discuss, one practical bounding exercise is to assume marginal costs are zero, so all social costs are simply the cost of funds. In that case, we need only multiply the $95.88 in program savings we estimate by the marginal cost of funds, which is usually calibrated at 0.3-0.5.

We discuss this in a few places in the revised paper. First, we added the following footnote to the model section discussing costs:

> *We note that, for a complete welfare analysis, this "cost" should be the social cost of the drug. This would generally be the marginal cost of production of the drug, not the price of the drug paid by Medicare or some other insurer. However, given that Medicare is a tax-financed program, we could also think of the "cost" as the marginal cost of funds required to finance the price paid by Medicare. For this section, we remain agnostic about the precise definition of the cost, and we come back to this question when we attempt to analyze welfare in Section 5.*

We have also modified our discussion of the welfare calculations in Section 5 as follows:

> *We have estimated the sum of the second and third terms to be \$86.*[a]
>
> [a] As noted in Section 1.2, this is not quite accurate for a complete welfare analysis because the second term is not the reduction in program costs but the reduction in social costs. We come back to this at the end of this section.

Finally, when walking through the comparison of lost consumer surplus to net savings, we have modified our discussion as follows:

> *A complete welfare analysis also requires us to map from net program savings to net social savings. When prior authorization reduces utilization, it reduces program costs by reducing the amount of money spent by payers (the insurer and the government). However, its effects on social costs come through two channels: (1) the social marginal cost of public funds needed to finance the insurance program; and (2) the marginal cost of producing the forgone drugs. Our data does not allow us to evaluate (2) here, so we treat these costs as if they are zero and consider our calibration here an excessively pessimistic measure of effects on social welfare. For (1), we use typical calibrations of the marginal cost of financing a dollar of public spending (\$0.30-\$0.50), and multiply them by the change in spending induced by prior authorization. This results in social savings from reduced drug spending of \$28-\$48. When compared to administrative costs of \$10, this results in net social financial savings of \$18-\$38. In the case of perfect screening, when the social marginal cost of public funds is high, prior authorization still improves social welfare, albeit very slightly. In other cases, prior authorization reduces social welfare, especially when screening is highly imperfect. In general, aside from the administrative costs, under this calibration, reducing drug consumption at all will generally reduce social welfare since we have assumed marginal cost of production is negligible.*

(R3.1b) *Even if prior authorization restrictions are welfare-improving ex-post, it is not clear that they are ex-ante. In an insurance market, evaluating any policy change would need to think about the trade-off between insurance and moral hazard. Is there a compelling reason to take the ex-post view? The authors should at least discuss how an ex-ante perspective might affect their conclusions, and the empirical objects they would need to estimate to perform that comparison.*

This is a good point that we had not previously considered. In our model, we assume no uncertainty about demand for drugs. You are absolutely correct that we could think about prior authorization as exposing beneficiaries to greater risk, since even those who don't, ex post, take a restricted drug have some ex ante probability of wanting it and now face greater barriers to getting it. We have mentioned this in a new footnote in Section 1.2:

> *We note that this assumes that there is no uncertainty about the drugs that a given beneficiary will need over the course of the next year. With uncertainty, there would be some "insurance value" to having drugs unrestricted, as with prior authorization restrictions the beneficiary would have a worse outcome in the "bad" state in which they need the restricted drug versus the outcome in the bad state without prior authorization. The removal of prior authorization thus plays a small role in equating marginal utilities in the good and bad states. This seems like a reasonable assumption here where most drug consumption is fairly persistent and predictable over time, suggesting little uncertainty regarding drug consumption and thus little insurance value. If, however, we assessed the uncertainty from "behind the veil of ignorance," insurance value could be larger (Hendren 2020).*

Unlike your other comment, we do not implement this in our welfare analysis because we have no obvious candidate for the appropriate coefficient of risk aversion to apply in this specific context.

(R3.2) *Counterfactual: The results in the paper consider a counterfactual where prior authorization restrictions are eliminated, but plan formularies and drug prices remain the same. Is there an actual policy or market intervention where this is plausible? If prior authorization were banned or simply not used, insurers would likely respond by (1) raising premiums, or (2) excluding more drugs. We also might expect drug prices (rebates) to be renegotiated. Explicitly considering insurers' choices of plan formularies and premia and bargaining between insurers and drug companies is way beyond the scope of this paper. But at the very least, some discussion is warranted. Are those margins of adjustment likely to be small? It would be helpful to understand the potential relevance of each margin of adjustment since it would matter for the effect on health care spending and the overall welfare impacts. For example, how often are the Medicare Part D minimum coverage requirements binding? Where they do bind, an insurer couldn't just exclude a drug currently subject to prior authorization.*

We agree that 1) this is not a plausible counterfactual; and 2) that there are many general equilibrium considerations, including the ones you highlight. We don't think that we have any firm sense of how important these considerations are, and they could indeed be quite important. The Editor has instructed us that he thinks this is too far afield of what is currently in the manuscript.

Our response, following his suggestion (E.8), is that we have significantly toned down the language that we use to describe conclusions from this and the below exercise, e.g. being much more cautious about using the words "inefficient" and "welfare" to describe our results.

(R3.3a) *Why not use the nested logit demand system to estimate the distribution of WTP in Section 6? For one thing, it would allow the authors to account for substitution to other drugs in*

the same class. It also seems like a natural robustness check on the constant elasticity demand assumption that is more consistent with the analysis of substitution patterns in Section 4.

There are a variety of trade-offs associated with using the logit versus our linear diff-in-diff approach for estimating WTP. However, per the instructions of the editor, in this resubmission we shifted everything we could to be based on linear models, including analyses of substitution *and* WTP, reserving the logit only for the simulations comparing the effects of prior auth on spending versus administrative costs. It is also the case that the current version of the demand model would not work for estimating WTP, as it is estimated off the drug choices of those randomized to plans, and these individuals, by definition, do not pay copays. There is no individual who pays copays who is randomized to plans, as you are only randomized if you are (1) defaulted into the LIS or (2) your prior plan was cancelled and you did not actively choose that plan initially. The people in (1) enter the LIS at the time they enter Medicare, so they do not experience a transition from copays (non-LIS) to copays (LIS). The people in (2) are a subset of the people from (1), as the only way to have not actively chosen your current plan is to have been defaulted into it, which only happens at Medicare entry. It could be feasible to use the logit to estimate WTP using the sample of individuals moving to the LIS, but this would again result in different samples, different models, etc., meaning there is no perfectly consistent solution here. Because of this, we opted to retain the reduced form demand estimation of WTP and shift it to be linear rather than constant semi-elasticity, as suggested by the editor.

All of that said, in the spirit of this comment, we now test the robustness of our WTP estimates to alternate functional forms for the demand curves - constant semi-elasticity of demand, linear demand, and class-specific demand. These results can be found in Table 10.

(R3.3b) *How do we think about the right "price" in Medicare Part D given the dynamic, non-linear nature of those contracts? A patient who is certain to reach the catastrophic coverage region (due to consumption of other drugs) faces a very low effective price, so actual copayments may not be the right price here. I worry this leads the authors to overestimate WTP.*

When we compute prices, we average over the spot price *actually faced* by patients. Therefore, if most patients who take the drug do so in the catastrophic region, the average price will be close to zero. While we are averaging over spot prices and a rational consumer would use their shadow price, there is enough evidence that consumers respond to spot prices that we feel comfortable doing this (Brot-Goldberg et al. 2017, Abaluck et al. 2018, Dalton et al. 2020). We acknowledge this in a new footnote in Section 5:

> *The "correct" price to use here is not immediately obvious, given the non-linear price schedule. A rational consumer would respond to the expected end-of-year price. However, there is substantial evidence that consumers also respond to 'spot' prices (Brot-*

> *Goldberg et al. 2017, Abaluck et al. 2018, Dalton et al. 2020), so we opt to use the actual spot prices paid by consumers.*

We also note that we've generally tried to downplay the use of the WTP estimates to evaluate welfare throughout Section 5 for this reason and due to other potential wedges between valuation and demand.

(R3.3c) *From section 2.1, it sounds like there is another group of LIS beneficiaries who pay some costs out-of-pocket. How is this group related to the group of patients used to estimate WTP in Section 6, who transitioned into the LIS program? Is there a way to use these partial-LIS patients to estimate demand (perhaps some of them transition to full LIS)? They may be more similar to the full-LIS population than patients who transition onto the program in terms of financial resources.*

The partial LIS are not randomly assigned in the way that the full LIS are, so we cannot use our research design to estimate their demand response.

(R3.MC1) *This may not directly impact the identification strategy, but it would be helpful to understand where the variation in prior authorization in a given market and year comes from. It's clear that there is a lot of variation in prior authorization restrictions across drug types, and for the same drug across plans within a market. Are these differences largely explained by which insurance company offers the plan (maybe some insurers tend to use prior authorization, others not)? Do specific drugs go on/off prior authorization over time for the same plan and market? Do insurers impose these restrictions on different drugs in different markets (for the same "plan")?*

This is a helpful thought, in line with similar comments by R1.1 and R2.1. What we did was take the residual variation in the use of prior authorization in assigned plan, net of drug-region-year fixed effects, and measure how much of its variation we can explain. We described this in Section 2.4:

> *This is not explained by some insurers being more prone to using prior authorization restrictions than others. Within drug-region-years, only 0.8% of residual variation in the use of prior authorization is explained by carriers. Even if we consider carrier-by-therapeutic class variation, this only explains 7.8% of residual variation.*

The drugs going on/off prior authorization over time within plans/regions should not explain any of our variation, since we always do analysis *within* regions and years. That said, in any given year, some plans may have prior auth on a drug while others don't simply because they turned prior auth on/off more quickly than other plans that then follow suit in the following years. But again, our variation is all within-market-year. We do not rely on variation over time, but any time variation in prior auth status within a drug across plans may be an explanation for the existence of within-market-year variation.

It would be challenging for us to think about the same plan in different markets, as this is not well-tracked in the Part D data. To the extent that we can track it (using contract and plan IDs), it appears that plans have uniform formularies across regions. If carriers use different plan IDs to represent the same plan with different formularies across regions, we would have a lot of difficulty matching them together; it would need to be done by hand, and there are many plans.

(R3.MC2) *What happens if a patient decides they want to pay for a restricted drug out-of-pocket? If the half of marginal patients who substitute to "no drug" are just buying it themselves, that would change the cost and welfare implications. Is there any evidence on how often this occurs? Would patients be penalized in any way (e.g. forgoing reimbursement for later use of those drugs) if they do?*

We have no way of tracking fully out-of-pocket purchasing. Our guess is that it's quite rare given that we study a low-income population, and the relevant drugs here are branded drugs which are likely to be quite expensive. However, if patients decided to pay out of pocket, they would not face any penalty, and the insurer would not be aware of this action (since a claim would not be filed).

(R3.MC3) *Prior authorization also involves delay, which may impose a cost on patients who do receive the drug separate from any administrative burden. Are there types of drugs and medical conditions where prescribing the drug is time-sensitive. We might then think of these costs as being significant, especially if the patients who do receive the drug have especially high values.*

The best example of a time-sensitive drug that we could come up with is oral anticoagulants, where just a day or two of not taking the drug can increase the risk of stroke significantly. Interestingly, in our case study around oral anticoagulants, we found that reductions in use of restricted drugs were completely offset by substitution to unrestricted oral anticoagulants, consistent with your hypothesis that for drugs that are really important we might expect to see less diversion to no drug. However, the editor did ask us to remove the case study, so we omit it from the revised draft.

(R3.MC4) *Natural dimensions along which to look at heterogeneity in prior authorizations' effects on spending are geography and insurer.*

We have now included heterogeneity analyses for service region (Appendix Figure A4) and insurance carrier (Appendix Figure A5). There is little carrier-level heterogeneity (which is reassuring, since our analyses with no plan FEs essentially assume carrier heterogeneity away on many dimensions). There is some regional heterogeneity but we don't think there is much to take away from the results.

Reviewer 4

(R4.1) *Summary evaluation: This is a very good paper. On the pro side, the big picture question is important and few papers have tried to answer it rigorously. The implementation in the paper is clear, careful, and comprehensive. The identifying variation for the main utilization effects is very clean. There are numerous convincing specification checks. On the con side, the LIS population studied is special in important ways. The paper bounces between different empirical specifications without sufficient explanation. And the exposition of the consumer surplus analysis is rushed.*
Broad takeaways: The population the authors consider is important but specific - a very subsidized no-copay program that's tightly linked to a less subsidized program (regular Part D) where copays ARE used to ration drugs. In fact, the magnitude of the cost savings in this LIS analysis here depends on prices negotiated between insurers and drug companies over both LIS and regular enrollees. So this paper tells us pretty convincingly how prior auth impacts LIS enrollees in the Medicare Part D program, but has less to say about how it impacts outcomes for regular Part D enrollees who also pay copays, Medicaid enrollees who face little cost-sharing but whose prices are determined differently, or employer-based insurance enrollees. Or how it would impact spending for LIS enrollees if the regular Part D program weren't there to temper prices. Prior auth is used in combination with different rationing tools and those combinations interact to impact which drugs are excluded, prices, and welfare in equilibrium. This is a limitation. One way to shed light on the role of regular Part D would be to examine how the use and effects of prior auth compare for plans with few vs. many LIS enrollees, if there is enough such variation. The authors could also discuss how prior auth varies across insurance settings and what that may mean for broad takeaways. One nice thing is that they offer a pretty general framework, which could be used to study other contexts.

Thank you for the kind words "on the pro side." We hope that this revision helps alleviate your "on the con side" concerns. In the revised manuscript, we follow the editor's instructions and try to make our empirical specifications more consistent throughout the paper. As per R3.MC4, we also estimate heterogeneous effects by carrier, and find little heterogeneity, which we think is informative about your "few vs. many LIS" suggestion. We also now discuss the similarities between our setting and others in a footnote (see also R1.5):

> *Both the forms and the sequence of events required for authorization for drugs, or for other services, are broadly similar for insurers in other U.S. insurance market segments.*

(R4.2) *Different empirical specifications: There are a lot of different empirical specifications, and the paper is not clear on why each is needed. That adds up to make the analysis seem more complicated than it is, and the paper less accessible to a general interest audience. If I'm reading the paper right, the demand model in section 4 relies on the same instrumental variable strategy and identifying variation as in section 3, but also factors in which drugs are substitutes for each other and in doing so enables counterfactual spending simulations. It also accounts for prior auth/exclusion of substitute drugs in a more straightforward way, and moreover is necessary to simulate D(0) for the welfare analysis. What's the unique benefit of section 3's more limited empirical specification, that justifies including both separately? What's so different about "The Effect of Authorization Restrictions on Drug Utilization" and "Substitution Patterns and Spending Effects"? There is yet another empirical specification in equation (3) in section 6. Why couldn't the authors just add price to the demand model from section 4 and estimate on the LIS-transition sample, maybe reweighting to better match the sample in section 4? Then those estimates could be used directly to get at welfare. There may be good reasons for these different empirical specifications, but the reasons aren't clear in the paper as written.*

This was an extraordinarily helpful comment that served as a bit of a touchstone as we revised. E.2 expresses this view quite strongly as well.

In response, we split the "Substitution Patterns and Spending Effects" into a substitution patterns part (folded into the main effects section) and a "counterfactual simulations"/total spending effects part (folded into a comparison against admin costs). In the main effects section, we now emphasize the use of simple linear models to estimate both effects on the focal drug and substitution effects.

We have retained the logit, but strip the discussion of it substantially, and emphasize its necessity only for estimating the effect of (counterfactually) removing *all* prior auth restrictions (which we see as the key policy counterfactual - banning prior auth), as well as providing an appropriate benchmark against which we can compare the change in admin costs (which we can really only calibrate for the full removal of prior auth rather than a one-off flipping of the prior auth switch for one specific drug, due to potential substitution between restricted drugs). See the new text below:

> *Importantly, simply scaling up our reduced form estimates may fail to account for interactions between restrictions on drugs. For instance, restricting drugs A and B is likely to have different effects than the sum of the effects of restricting each drug individually. This is because restricting A may induce substitution to B (and vice versa). These interactions may be important for understanding the full extent of the effects of allowing prior authorization as an institution. Further, turning restrictions off for one drug at a time (the effect captured in our reduced form approach) may induce substitution to another restricted drug, thus not reducing administrative costs. As discussed earlier, we lack the variation to cleanly identify such rich substitution*

> *patterns, making this empirically difficult without imposing additional structure.*
>
> *Thus, we proceed by explicitly modeling demand for drugs using a standard micro-founded model of demand for restricted and unrestricted drugs (and their substitutes) based on utility maximization. We then use the estimated model to simulate removing all prior authorization restrictions and measure the effects on utilization, spending, and administrative costs. Such an approach requires us to make additional data restrictions and modeling assumptions relative to our reduced-form approach, but we show that these choices have relatively minimal effects via comparisons to our reduced-form estimates.*

One reason why we cannot take the strategy you (and R3.3a) suggest of estimating the demand model on those who transition into the LIS is that the transitioners do *not* face random assignment as they transition, since they are already actively enrolled in a plan. Similarly, those that do face random assignment by definition never face out-of-pocket costs, thus making it impossible for us to estimate WTP using that group of beneficiaries.

(R4.3) *Consumer surplus: The authors are admirably transparent about the pros and cons of revealed preference for measuring welfare. And the authors clearly have smart and nuanced things to say about welfare in this setting. But the exposition in that section is quite brief given how complicated it is. WHY might the set of individuals rationed out first under prior auth be different from the set of individuals rationed out first under positive prices? What would that look like in the model in section 1.2?*

This is a good point. We now elaborate on the reasons for different screening under prior auth versus copayments in Section 5:

> *However, responsiveness to price may not be perfectly reflective of responsiveness to prior authorization. For instance, physicians play a major role in the prior authorization process, and their beliefs about the value of a given drug for a given beneficiary may differ from the patient's own values. Similarly, physicians may differ in their administrative capacity to deal with the authorization process (Gandhi and Shi, 2024), and patient matching to different physicians need not be related to their WTP for specific restricted drugs. Therefore, which beneficiaries get screened out by prior authorization may be completely orthogonal to their WTP.*

(R4.4) *Regarding implementation, I thought as I read the introduction that the drug demand analysis in section 4 was going to be used to calculate WTP for drugs rationed out under prior auth. Instead the authors use a sample of enrollees transitioning LIS status infer an average price semi-elasticity for all drugs, then trace out the demand curve using the price elasticity and*

demand at zero price from section 4, then calculate lost surplus implied by that demand curve under two assumptions about where the individuals rationed out under prior auth are. On the upside, it's great that the observed prior auth seems efficient even under the pretty conservative assumption that the implicit targeting that prior auth entails is at least better than random. On the downside, it's yet another empirical model, with a single demand curve for all drugs that feels incongruous with the heterogeneity documented elsewhere. It seems simpler to just use the demand model in section 4 for all purposes in this paper. If that's not feasible, the authors should say why, and touch on whether the consumer surplus results are robust to alternative parametric assumptions. It probably doesn't matter for the efficient rationing assumption, but may for the random rationing assumption. Einav Finkelstein and Williams (AEJPol 2016) section III(C) would be a nice model for this discussion. (Of course, using the demand model from section 4 would be a parametric approach as well, but at least it'd all be internally consistent.)

This is a good point. There are a variety of trade-offs associated with using the logit versus our linear diff-in-diff approach for estimating WTP. However, per the instructions of the editor, in this resubmission we shifted everything we could to be based on linear models, including analyses of substitution *and* WTP, reserving the logit only for the simulations comparing the effects of prior auth on spending versus administrative costs. It is also the case that the current version of the demand model would not work for estimating WTP, as it is estimated off the drug choices of those randomized to plans, and these individuals, by definition, do not pay copays. There is no individual who pays copays who is randomized to plans, as you are only randomized if you are (1) defaulted into the LIS or (2) your prior plan was cancelled and you did not actively choose that plan initially. The people in (1) enter the LIS at the time they enter Medicare, so they do not experience a transition from copays (non-LIS) to copays (LIS). The people in (2) are a subset of the people from (1), as the only way to have not actively chosen your current plan is to have been defaulted into it, which only happens at Medicare entry. It could be feasible to use the logit to estimate WTP using the sample of individuals moving to the LIS, but this would again result in different samples, different models, etc., meaning there is no perfectly consistent solution here. Because of this, we opted to retain the reduced-form demand estimation of WTP and shift it to be linear rather than constant semi-elasticity, as suggested by the editor.

(R4.5) *It would be nice to see more discussion of the NICE results, since they're at least intended to capture differential health benefits of drugs relative to their substitutes. One could see another similar paper using the NICE classifications or the clinical information underlying them to get at welfare. The authors must not think that's a good idea. Why?*

We initially hoped to use the NICE data to do the kind of analysis suggested by the reviewer. However, the data have important limitations. First, the public NICE data only contains three categories, which

are based on cost effectiveness to the English National Health Service (i.e. the ratio of differential health benefits to procurement cost). We cannot observe either the estimated health benefits of an evaluated drug relative to its comparators nor the cost of the drug to the NHS. Procurement costs can differ substantially between the UK and the US. Second, NICE reports are only available for a subset of the drugs in our sample. We agree with the reviewer that trying to understand the likely health losses from foregone drug consumption is an important area for future research.

(R4.6) *In Table 1, how are these sample Medicare enrollees under 65 on average? The average age in Einav Finkelstein and Schrimpf 2015 is 71.*

In contrast to those authors, we do not exclude Medicare enrollees who qualify due to disability. We now make this more clear in a footnote in Section 2.3:

> *Note that, in our analysis, we retain those who qualify for Medicare due to disability rather than old age, lowering the average age of our population.*

We decided to include this population because most full LIS recipients are dual-eligible, and a large share of dual-eligible beneficiaries are under 65.

(R4.7) *Why are excluded drugs in the choice set in sections 3 and 4?*

As we note on in Section 2.2, our measure of whether a drug is excluded is whether it is not listed on the plan's formulary. We now include a footnote mentioning:

> *Additionally, note that our definition of exclusion is given by non-inclusion, so it is possible that some drugs we designate as being excluded are covered but this coverage is not reported by the plan to CMS. Our claims data includes covered claims for drugs we designate as excluded, which may either reflect mis-classification or insurer-granted exceptions.*

We also note that exclusion is not absolute, in that beneficiaries can appeal for coverage for excluded drugs. Instead, we think exclusion should be seen as an extreme version of prior authorization, with a much less clear path to coverage and potentially observed with measurement error.

(R4.8) *What is the weight formula at the bottom of page 26 doing?*

See our response to E.5.